Journal of the American College of Cardiology Vol. 56, No. 23, 2010

Published by Elsevier Inc.
ISSN 0735-1097/$36.00
doi:10.1016/j.jacc.2010.10.004

Clinical Guidelines and Performance Measures

Responsible Guidance and Accountability

Frederick A. Masoudi, MD, MSPH,* Eric D. Peterson, MD, MPH,† Jeffrey L. Anderson, MD,‡ Robert O. Bonow, MD,§ Alice K. Jacobs, MD‖

Denver, Colorado; Durham, North Carolina; Salt Lake City, Utah; Chicago, Illinois; and Boston, Massachusetts

Ghali et al. (1) raise important points about guidelines and performance measures. Although they focus on heart failure, their comments are relevant beyond a single clinical condition. We agree with the goal of personalizing patient care to the extent possible, but we do not believe that guidelines and performance measures necessarily conflict with this important goal. As members of the American College of Cardiology Foundation (ACCF) and American Heart Association (AHA) Task Force on Practice Guidelines (TFPG) and Task Force on Performance Measures (TFPM), we welcome the opportunity to discuss our perspective on these issues.

Clinical Practice Guidelines

Since the early 1980s, the ACCF and the AHA have engaged in a joint effort to develop clinical practice guidelines, consistent with the mission of both organizations to improve the quality of care and outcomes for patients with cardiovascular disease and fostered by evidence of substantial variation in rates of adherence to evidence-based therapies. Each guideline represents a summary and synthesis of the available evidence by a writing committee, which is vetted through an exhaustive process of peer review and approval by the TFPG, the Board of Trustees of the ACCF, and the Science Advisory Coordinating Committee of the AHA, in addition to the governing bodies of relevant partnering organizations. Therefore, it is not surprising that practice according to ACCF and AHA guidelines recommendations is associated with improved patient outcomes (2,3).

The methods used to collect, evaluate and classify the evidence adopted by the TFPG are widely known (4,5). Briefly, the recommendations are classified according to whether a test, procedure, treatment, or strategy is useful and effective considering the size of the treatment effect (Class I, IIa, IIb, and III), relying where possible on randomized clinical trials (RCTs) and according to an estimate of the certainty (or precision) of the treatment effect (Level of Evidence: A, B, or C). Recommendations based on limited data, case reports, consensus opinion, or standard of care (or on sound clinical judgment where it is not feasible to study the issue) are considered Level of Evidence: C.

We agree with the concern of Ghali et al. (1) that we need to expand our evidence base to more definitively guide treatment choices across the spectrum of patient types and clinical scenarios faced in routine practice. The Level of Evidence: C recommendations, however, reflect the TFPG's desire to bring together expert consensus opinion to more fully advise clinicians in common clinical situations for which limited or no evidence yet exists. Moreover, it is precisely when evidence is lacking or conflicting that clinicians frequently need and request guidance from experts.

We also agree that RCTs, while providing some of the most rigorous evidence for therapies in selected populations, have important limitations. This view is incorporated into the TFPG methodology. Each RCT that is reviewed,

From the *Denver Health Medical Center, Denver, Colorado; †Duke Clinical Research Institute, Durham, North Carolina; ‡Intermountain Health Care, Salt Lake City, Utah; §Northwestern University, Chicago, Illinois; and the ‖Boston University Medical Center, Boston, Massachusetts. Dr. Masoudi is the chair of the ACCF/AHA Task Force on Performance Measures; is an Associate Professor of Medicine at the Denver Health Medical Center Division of Cardiology; has served as a consultant for Axio Research; and has received financial benefit from the American College of Cardiology, American Heart Association, the Oklahoma Foundation for Medical Quality, and the Massachusetts Medical Society. Dr. Peterson is the incoming chair for the ACCF/AHA Task Force on Performance Measure (term begins in January 2011); is a professor of medicine and director of cardiovascular outcomes for the Duke Clinical Research Institute; is a consultant for AstraZeneca and Tethysbio; and has done personal research for Bristol-Myers Squibb/Sanofi, Eli Lilly, Johnson & Johnson, and Merck and Schering. Dr. Anderson is the incoming chair of the ACCF/AHA Task Force on Practice Guidelines (term begins in January 2012); is an Associate Chief of Cardiology for the Intermountain Medical Center; has done research for the AstraZeneca, FUTURE-OASIS-CURRENT study, TIMI-48,-51, and -52 studies, and Toshiba; has received financial benefit from the COAG study, CORAL study, CoumaGenil study, and the GIFT study; and was an expert witness in 2010 for the defense for the management of cardiopulmonary arrest post-op and stroke after ablation for atrial fibrillation. Dr. Bonow is the past chair of the ACCF/AHA Task Force on Performance Measures; is the Chief of the Division of Cardiology, for Northwestern University Feinberg School of Medicine; and is a consultant for Edwards Lifesciences. Dr. Jacobs is the chair of the ACCF/AHA Task Force on Practice Guidelines; is a professor of medicine and director of the cardiac catheterization lab for Boston University Medical Center; and has done personal research for Abbott Vascular, Abiomed, Accumetrics, Cardiovascular Research Foundation, Harvard Clinical Research Institute, and the TIMI study group. This paper is being simultaneously copublished in the *Journal of Cardiac Failure*.

Manuscript received September 27, 2010; accepted October 4, 2010.

especially if it informs a recommendation, is evaluated for quality and overall consistency with existing evidence base. Often the limitations of a study are reflected in the specific language of the recommendation or in the accompanying text (such as the recommendation for implantable cardioverter-defibrillator placement for patients with left ventricular systolic dysfunction, which recommends that implantation only be considered in patients with a "reasonable expectation of survival with a good functional status for more than 1 year") (6). We also recognize the trade-off between internal validity (of a RCT) and external validity or generalizability (of a registry or clinical practice experience), but in view of the known limitations of (even pre-specified) subgroup analyses (7), specific subgroups are usually not culled out of any recommendation absent robust data to support this approach. When available, the committees also consider whether trial results from select population are also supported by comparative effectiveness evaluations in broader observational registries. Where evidence exists, specific sections in the guideline will address the elderly, patients with diabetes, women, and others. Yet in general, patients are typically more similar than they are different. When studied in sufficient numbers of patients, true differences in treatment effectiveness or safety are rare. As a result, the TFPG appropriately does not recommend differential therapy for women or the elderly simply on the basis of hypothesis-generating subgroup analyses from RCTs.

Given these considerations, the preamble to every guideline states that these statements are intended as general guides and that patient care decisions are ultimately to be made by the treating clinician in view of all of the circumstances presented by a specific patient. The preamble also states that prescribed courses of treatment in accordance with recommendations are effective only if they are followed. Lack of patient adherence or a failure of clinicians to deliver or monitor an indicated therapy may adversely affect outcomes. Therefore, physicians and patients must work together to promote optimal and safe treatment regimens and lifestyle changes.

Finally, we realize that the guideline process is not perfect. Rather, it is iterative, and currently, changes to the evidence review, a specific evidence scoring system, the addition of Bayesian analysis, and shorter formats are being evaluated as enhancements to the process. These innovations will be described in an upcoming publication. We believe that these changes will improve the extent to which guidelines reflect existing evidence and facilitate translation of guidelines in practice.

Performance Measures

With respect to performance measures, we agree that not all guideline recommendations are appropriate for performance measurement. Indeed, as Ghali et al. (1) point out, although the heart failure guideline contains dozens of Class I recommendations, a select few of these processes of care have been incorporated into performance measures. The strongest evidence base, while necessary for the development of performance measures, is by no means sufficient. As delineated in detailed methodology developed by the TFPM, performance measures must also be valid, reliable, actionable, and measurable, and they must address a demonstrable gap in care (8). The TFPM has also developed methodology beyond processes of care, delineating detailed "best practices" for the development of measures of outcomes (9) and efficiency (10) and for composite measures (11). Thus, although the strongest guideline recommendations define candidates for performance measurement, only those processes that meet a number of other criteria will ultimately be considered appropriate for the purposes of accountability after the application of an explicit and rigorous methodology.

The TFPM also has supported the use of clinical exclusions in all process measures as a means of avoiding the potentially hazardous overuse of therapies in patients with contraindications (8,12). Specifically, for all process measures, explicit clinician documentation of a reason for not providing a therapy results in an exclusion, thus removing any incentive to provide therapies in patients who are not ideal candidates. Although payers and consumer groups do not necessarily share the perspective that these exclusions are necessary, the TFPM considers this characteristic as fundamental to the validity and acceptability of measures of processes of care (8,12). The assertion that performance measures diminish the vital role of physician judgment and insight into patient management does not account for this important methodological precept of ACCF and AHA performance measure development. ACCF and AHA performance measures are in fact designed to encourage clinicians to consider whether a therapy is indicated and not contraindicated in each patient, functioning as simultaneous stimuli to reduce both underuse and overuse of therapies.

Additionally, the TFPM strongly advocates the testing of its measures in practice to further assess their validity and reliability, to determine the burden of data collection, and to identify unintended consequences of measure implementation. The capacity to perform such testing has been conducted both in collaboration with outside organizations (e.g., the Centers for Medicare and Medicaid Services) and with the registry programs of the ACCF and AHA. Ultimately, this testing is necessary to determine the impact of measures on practice and patient outcomes.

Finally, as individuals actively involved in clinical practice, we are acutely aware of the frustrations and anxiety that performance measurement can engender. However, we are also cognizant that U.S. health care has moved beyond the question of whether performance should be measured; the focus now is how measurement will be conducted. Faced with the decision to shun performance measurement or to take an active role in it, the ACCF and AHA, in our opinion wisely, chose to engage. Because of this decision,

the ACCF and AHA have been able to play an important role at the national level in advocating that the measurement of quality in cardiovascular disease and stroke adheres to rigorous methodology. As professionals, we believe we have an obligation to define best practices on the basis of the evidence, to measure how consistently these practices are being used in our patients, and to stimulate future quality improvement efforts that promote safe and effective care and maximally improve patient outcomes.

Industry, Guidelines, and Performance Measures

The ACCF and AHA share the concern voiced by Ghali et al. (1) about the potential conflicts of interest from industry in the development of guideline and performance measures. Recognizing this, the ACCF and AHA have an explicit, and relatively stringent, policy on relationships with industry and other entities for writing group membership, including restrictions on the number of writing group members who can have relevant conflicts and restricting the writing and voting on specific sections of the documents to those individuals without relationships relevant to these sections (13).

Conclusions

The ACCF and AHA TFPG and TFPM are aware of the gap between "what we truly know, what we think we know, and what we would like to know." This gap is the precise reason for the characterization of the level of evidence in guideline recommendations, for the approach to writing recommendations, for the reluctance to generate recommendations concerning subgroups without adequate supporting evidence, and for the methodical process to determine which Class I guideline recommendations qualify for consideration of performance measures. We believe that guidelines and performance measures generated by the ACCF and AHA provide clinicians with methodologically rigorous and practical tools to support their practice. We are committed to continually evolving and updating both the development process as well as the documents themselves.

Reprint requests and correspondence: Dr. Frederick A. Masoudi, Division of Cardiology MC0960, Denver Health Medical Center, 777 Bannock Street, Denver, Colorado 80204. E-mail: fred.masoudi@ucdenver.edu.

REFERENCES

1. Ghali JK, Massie BM, Mann DL, Rich MW. Heart failure guidelines, performance measures, and the practice of medicine: mind the gap. J Am Coll Cardiol 2010;56:2077–80.
2. Fox KA, Steg PG, Eagle KA, et al. Decline in rates of death and heart failure in acute coronary syndromes, 1999–2006. JAMA 2007;297:1892–900.
3. Peterson ED, Roe MT, Mulgund J, et al. Association between hospital process performance and outcomes among patients with acute coronary syndromes. JAMA 2006;295:1912–20.
4. Methodology manual and policies from the ACCF/AHA Task Force on Practice Guidelines. Available at: http://www.cardiosource.org/Science-And-Quality/Practice-Guidelines-and-Quality-Standards/Methodology.aspx. Accessed September 25, 2010.
5. Gibbons RJ, Smith S, Antman E. American College of Cardiology/American Heart Association clinical practice guidelines: part I: where do they come from? Circulation 2003;107:2979–86.
6. Jessup M, Abraham WT, Casey DE, et al. 2009 focused update: ACCF/AHA guidelines for the diagnosis and management of heart failure in adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines: developed in collaboration with the International Society for Heart and Lung Transplantation. Circulation 2009;119:1977–2016.
7. Kaul S, Diamond GA. Trial and error. How to avoid commonly encountered limitations of published clinical trials. J Am Coll Cardiol 2010;55:415–27.
8. Spertus JA, Eagle KA, Krumholz HM, et al. American College of Cardiology and American Heart Association methodology for the selection and creation of performance measures for quantifying the quality of cardiovascular care. J Am Coll Cardiol 2005;45:1147–56.
9. Krumholz HM, Brindis RG, Brush JE, et al. Standards for statistical models used for public reporting of health outcomes: an American Heart Association scientific statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. Circulation 2006;113:456–62.
10. Krumholz HM, Keenan PS, Brush JE Jr., et al. Standards for measures used for public reporting of efficiency in health care: a scientific statement from the American Heart Association Interdisciplinary Council on Quality of Care and Outcomes research and the American College of Cardiology Foundation. J Am Coll Cardiol 2008;52:1518–26.
11. Peterson ED, DeLong ER, Masoudi FA, et al. ACCF/AHA 2010 position statement on composite measures for healthcare performance assessment: a report of American College of Cardiology Foundation/American Heart Association Task Force on Performance Measures (Writing Committee to Develop a Position Statement on Composite Measures). J Am Coll Cardiol 2010;55:1755–66.
12. Spertus JA, Bonow RO, Chan P, et al. ACCF/AHA new insights into the methodology of performance measurement: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Performance Measures. J Am Coll Cardiol 2010. In press.
13. ACCF relationship with industry and other entities: policies and procedures for the development of guidelines and other clinical documents. Available at: http://www.cardiosource.org/Science-And-Quality/Practice-Guidelines-and-Quality-Standards/Relationships-With-Industry-Policy.aspx. Accessed September 23, 2010.

Key Words: guidelines ■ performance measures ■ practice of medicine.